A Novel Privacy-Preserving Socio-technical Platform for Detecting Cyber Abuse (*Extended Abstract*)

Sriram Chellappan(✉) and Nathan Fisk

Department of Computer Science and Engineering and College of Education, University of South Florida, Tampa, USA
{sriramc,fisk}@usf.edu

Abstract. In this abstract, we present perspectives on on-going work by the authors in creating a massive scale digital platform that enables youth to share content that suits their own privacy expectations, while still contributing to cyber-abuse research.

1 Design of Our Data Collection Platform

Currently, research in cyber-abuse suffers from two problems: (a) linguistic differences between younger people (i.e., victims) and adults (that flag abuse) in cyber space [1] and (b) IRB regulations that are strict in protecting vulnerable subjects (e.g., children) in research [2]. As a result, existing research in cyber-abuse lacks robust ground truth datasets, that incorporates victim perspectives.

Figure 1 illustrates our platform that aims to correct the above trend. We design a smart-phone app that allows users (in this case younger people) to download the app after their parents/guardians consent. Users will then be asked a series of questions on the app related to situations of discomfort they have faced in the past, along with a timeline of related event(s). Subsequently, meta-data of communication logs that are in the vicinity of the event(s) described are searched and retrieved within the device, and displayed to the user. The meta-data logs in our current prototype are only (a) the exact times of communication; (b) whether the communication was sent or received; (c) the sizes of the content of communication (i.e., sizes of textual or imagery or video content); (d) whether the communication was bi-directional or within a group; and (e) results of sentiment analysis algorithms applied on textual messages. For our prototype, only SMS logs are processed, but the system can be easily expanded to other content with right permissions. All user identities are anonymized. These meta-data logs (which are highly privacy preserving) are exported upon user assent to a secure cloud, where a domain expert will process user entered responses and follow up

This work was supported in part by US National Science Foundation. Findings and conclusions are those of the authors alone, and not necessarily those of the funding agency.

A. Tagarelli and H. Tong (Eds.): CSoNet 2019, LNCS 11917, pp. 307–308, 2019.
https://doi.org/10.1007/978-3-030-34980-6_33

with a subject on a need basis via the app only. Note though all user responses are filtered if any identifiable information like names, phone numbers, emails etc. are provided before uploading to the cloud.

2 Challenges

We are currently demoing our prototype, and surveying many stakeholders - young people, parents, counselors, tech developers and privacy/law experts also to assess our system from multiple angles including usability, efficacy, privacy-preservation and maintenance cost. We are already getting interesting results. While there is near unanimous agreement on the fact that perspectives of victims (in this case young people) are considered for abuse detection, there is some push back from the tech community in terms of data sharing (even meta-data). There were concerns raised on long-term privacy implications, especially on the power of longitudinal meta-data to identify individuals. Furthermore, with meta-alone, there are interesting algorithmic challenges ahead for us on identifying and modeling network topologies, especially under churn.

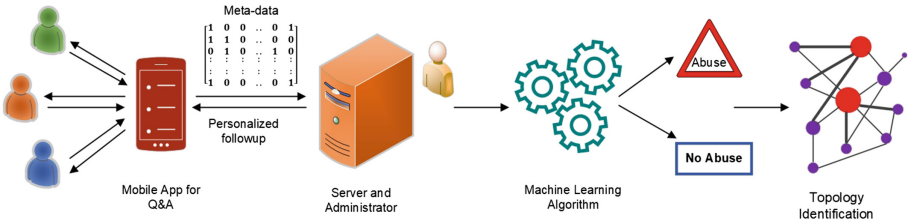

Fig. 1. Workflow of our platform

3 Insights for Future

Should we resolve all challenges (technical and social) satisfactorily, we will be piloting our system soon. With larger scale participation and massive scale meta-data, future work includes design of algorithms to (a) flag abuse at early stages; (b) classify the type of abuse; (c) detect changes in patterns of abuse at the individual level; (d) identify group dynamics during abuse; (e) model the evolution of the generated social network toplogies and so much more. Ensuring privacy, while simultaneously keeping the end users engaged on this important global problem today are the most significant aspects of our research.

References

1. Selwyn, N.: The digital native-myth and reality. In: Aslib Proceedings, vol. 61, pp. 364–379. Emerald Group Publishing Limited (2009)
2. Glantz, L.H.: Conducting research with children: legal and ethical issues. J. Am. Acad. Child Adolesc. Psychiatry **35**(20), 1283–1291 (1996)